

## **RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical  
Information Center (STIC) no errors detected.**

Application Serial Number: 10/777,144

Source: \_\_\_\_\_

Date Processed by STIC: \_\_\_\_\_

# ***ENTERED***



IFWO

## RAW SEQUENCE LISTING

DATE: 10/18/2004

PATENT APPLICATION: US/10/777,144

TIME: 09:18:43

Input Set : N:\CrF3\RULE60\10777144.raw.txt

Output Set: N:\CRF4\10182004\J777144.raw

## SEQUENCE LISTING

3 (1) GENERAL INFORMATION:

5 (i) APPLICANT: Schlessinger, Joseph

6 Sap, Jan M.

8 (ii) TITLE OF INVENTION: NOVEL RECEPTOR-TYPE PHOSPHOTYROSINE

9 PHOSPHATASE-ALPHA

11 (iii) NUMBER OF SEQUENCES: 14

13 (iv) CORRESPONDENCE ADDRESS:

14 (A) ADDRESSEE: PENNIE & EDMONDS

15 (B) STREET: 1155 AVENUE OF THE AMERICAS

16 (C) CITY: NEW YORK

17 (D) STATE: NEW YORK

18 (E) COUNTRY: U.S.A.

19 (F) ZIP: 10036

21 (v) COMPUTER READABLE FORM:

22 (A) MEDIUM TYPE: Floppy disk

23 (B) COMPUTER: IBM PC compatible

24 (C) OPERATING SYSTEM: PC-DOS/MS-DOS

25 (D) SOFTWARE: PatentIn Release #1.0, Version #1.25

27 (vi) CURRENT APPLICATION DATA:

C--> 28 (A) APPLICATION NUMBER: US/10/777,144

C--> 29 (B) FILING DATE: 13-Feb-2004

30 (C) CLASSIFICATION:

32 (vii) PRIOR APPLICATION DATA:

W--> 33 (A) APPLICATION NUMBER: US/09/280,597

34 (B) FILING DATE: 29-MARCH-1999

W--> 36 (A) APPLICATION NUMBER: US 08/015,985

37 (B) FILING DATE: 10-FEB-1993

39 (viii) ATTORNEY/AGENT INFORMATION:

40 (A) NAME: Coruzzi, Laura A.

41 (B) REGISTRATION NUMBER: 30,742

42 (C) REFERENCE/DOCKET NUMBER: 7683-020

44 (ix) TELECOMMUNICATION INFORMATION:

45 (A) TELEPHONE: (212) 790-9090

46 (B) TELEFAX: (212) 869-9741/8864

47 (C) TELEX: 66141 PENNIE

51 (2) INFORMATION FOR SEQ ID NO: 1:

53 (i) SEQUENCE CHARACTERISTICS:

54 (A) LENGTH: 802 amino acids

55 (B) TYPE: amino acid

56 (D) TOPOLOGY: linear

58 (ii) MOLECULE TYPE: protein

60 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

## RAW SEQUENCE LISTING

DATE: 10/18/2004

PATENT APPLICATION: US/10/777,144

TIME: 09:18:43

Input Set : N:\Crf3\RULE60\10777144.raw.txt

Output Set: N:\CRF4\10182004\J777144.raw

```

62 Met Asp Ser Trp Phe Ile Leu Val Leu Leu Gly Ser Gly Leu Ile Cys
63   1           5           10           15
65 Val Ser Ala Asn Asn Ala Thr Thr Val Ala Pro Ser Val Gly Ile Thr
66           20           25           30
68 Arg Leu Ile Asn Ser Ser Thr Ala Glu Pro Val Lys Glu Glu Ala Lys
69           35           40           45
71 Thr Ser Asn Pro Thr Ser Ser Leu Thr Ser Leu Ser Val Ala Pro Thr
72           50           55           60
74 Phe Ser Pro Asn Ile Thr Leu Gly Pro Thr Tyr Leu Thr Thr Val Asn
75 65           70           75           80
77 Ser Ser Asp Ser Asp Asn Gly Thr Thr Arg Thr Ala Ser Thr Asn Ser
78           85           90           95
80 Ile Gly Ile Thr Ile Ser Pro Asn Gly Thr Trp Leu Pro Asp Asn Gln
81           100          105          110
83 Phe Thr Asp Ala Arg Thr Glu Pro Trp Glu Gly Asn Ser Ser Thr Ala
84           115          120          125
86 Ala Thr Thr Pro Glu Thr Phe Pro Pro Ser Gly Asn Ser Asp Ser Lys
87           130          135          140
89 Asp Arg Arg Asp Glu Thr Pro Ile Ile Ala Val Met Val Ala Leu Ser
90 145          150          155          160
92 Ser Leu Leu Val Ile Val Phe Ile Ile Ile Val Leu Tyr Met Leu Arg
93           165          170          175
95 Phe Lys Lys Tyr Lys Gln Ala Gly Ser His Ser Asn Ser Phe Arg Leu
96           180          185          190
98 Ser Asn Gly Arg Thr Glu Asp Val Glu Pro Gln Ser Val Pro Leu Leu
99           195          200          205
101 Ala Arg Ser Pro Ser Thr Asn Arg Lys Tyr Pro Pro Leu Pro Val Asp
102          210          215          220
104 Lys Leu Glu Glu Glu Ile Asn Arg Arg Met Ala Asp Asp Asn Lys Leu
105 225          230          235          240
107 Phe Arg Glu Glu Phe Asn Ala Leu Pro Ala Cys Pro Ile Gln Ala Thr
108           245          250          255
110 Cys Glu Ala Ala Ser Lys Glu Glu Asn Lys Glu Lys Asn Arg Tyr Val
111           260          265          270
113 Asn Ile Leu Pro Tyr Asp His Ser Arg Val His Leu Thr Pro Val Glu
114           275          280          285
116 Gly Val Pro Asp Ser Asp Tyr Ile Asn Ala Ser Phe Ile Asn Gly Tyr
117          290          295          300
119 Gln Glu Lys Asn Lys Phe Ile Ala Ala Gln Gly Pro Lys Glu Glu Thr
120 305          310          315          320
122 Val Asn Asp Phe Trp Arg Met Ile Trp Glu Gln Asn Thr Ala Thr Ile
123           325          330          335
125 Val Met Val Thr Asn Leu Lys Glu Arg Lys Glu Cys Lys Cys Ala Gln
126           340          345          350
128 Tyr Trp Pro Asp Gln Gly Cys Trp Thr Tyr Gly Asn Ile Arg Val Ser
129           355          360          365
131 Val Glu Asp Val Thr Val Leu Val Asp Tyr Thr Val Arg Lys Phe Cys
132          370          375          380
134 Ile Gln Gln Val Gly Asp Met Thr Asn Arg Lys Pro Gln Arg Leu Ile

```

## RAW SEQUENCE LISTING

DATE: 10/18/2004

PATENT APPLICATION: US/10/777,144

TIME: 09:18:43

Input Set : N:\Crf3\RULE60\10777144.raw.txt

Output Set: N:\CRF4\10182004\J777144.raw

```

135 385          390          395          400
137 Thr Gln Phe His Phe Thr Ser Trp Pro Asp Phe Gly Val Pro Phe Thr
138          405          410          415
140 Pro Ile Gly Met Leu Lys Phe Leu Lys Lys Val Lys Ala Cys Asn Pro
141          420          425          430
143 Gln Tyr Ala Gly Ala Ile Val Val His Cys Ser Ala Gly Val Gly Arg
144          435          440          445
146 Thr Gly Thr Phe Val Val Ile Asp Ala Met Leu Asp Met Met His Thr
147          450          455          460
149 Glu Arg Lys Val Asp Val Tyr Gly Phe Val Ser Arg Ile Arg Ala Gln
150 465          470          475          480
152 Arg Cys Gln Met Val Gln Thr Asp Met Gln Tyr Val Phe Ile Tyr Gln
153          485          490          495
155 Ala Leu Leu Glu His Tyr Leu Tyr Gly Asp Thr Glu Leu Glu Val Thr
156          500          505          510
158 Ser Leu Glu Thr His Leu Gln Lys Ile Tyr Asn Lys Ile Pro Gly Thr
159          515          520          525
161 Ser Asn Asn Gly Leu Glu Glu Phe Lys Lys Leu Thr Ser Ile Lys
162          530          535          540
164 Ile Gln Asn Asp Lys Met Arg Thr Gly Asn Leu Pro Ala Asn Met Lys
165 545          550          555          560
167 Lys Asn Arg Val Leu Gln Ile Ile Pro Tyr Glu Phe Asn Arg Val Ile
168          565          570          575
170 Ile Pro Val Lys Arg Gly Glu Glu Asn Thr Asp Tyr Val Asn Ala Ser
171          580          585          590
173 Phe Ile Asp Gly Tyr Arg Gln Lys Asp Ser Tyr Ile Ala Ser Gln Gly
174          595          600          605
176 Pro Leu Leu His Thr Ile Glu Asp Phe Trp Arg Met Ile Trp Glu Trp
177          610          615          620
179 Lys Ser Cys Ser Ile Val Met Leu Thr Glu Leu Glu Arg Gly Gln
180 625          630          635          640
182 Glu Lys Cys Ala Gln Tyr Trp Pro Ser Asp Gly Leu Val Ser Tyr Gly
183          645          650          655
185 Asp Ile Thr Val Glu Leu Lys Lys Glu Glu Glu Cys Glu Ser Tyr Thr
186          660          665          670
188 Val Arg Asp Leu Leu Val Thr Asn Thr Arg Glu Asn Lys Ser Arg Gln
189          675          680          685
191 Ile Arg Gln Phe His Phe His Gly Trp Pro Glu Val Gly Ile Pro Ser
192          690          695          700
194 Asp Gly Lys Gly Met Ile Ser Ile Ile Ala Ala Val Gln Lys Gln Gln
195 705          710          715          720
197 Gln Gln Ser Gly Asn His Pro Ile Thr Val His Cys Ser Ala Gly Ala
198          725          730          735
200 Gly Arg Thr Gly Thr Phe Cys Ala Leu Ser Thr Val Leu Glu Arg Val
201          740          745          750
203 Lys Ala Glu Gly Ile Leu Asp Val Phe Gln Thr Val Lys Ser Leu Arg
204          755          760          765
206 Leu Gln Arg Pro His Met Val Gln Thr Leu Glu Gln Tyr Glu Phe Cys
207          770          775          780

```

## RAW SEQUENCE LISTING

DATE: 10/18/2004

PATENT APPLICATION: US/10/777,144

TIME: 09:18:43

Input Set : N:\Crif3\RULE60\10777144.raw.txt

Output Set: N:\CRF4\10182004\J777144.raw

```

209 Tyr Lys Val Val Gln Glu Tyr Ile Asp Ala Phe Ser Asp Tyr Ala Asn
210 785                      790                      795                      800
212 Phe Lys
215 (2) INFORMATION FOR SEQ ID NO: 2:
217     (i) SEQUENCE CHARACTERISTICS:
218         (A) LENGTH: 2409 base pairs
219         (B) TYPE: nucleic acid
220         (C) STRANDEDNESS: double
221         (D) TOPOLOGY: unknown
223     (ii) MOLECULE TYPE: cDNA
225     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
227 ATGGATTCCCT GGTTCATTCT TGTTCCTGCTC GGCAGTGGTC TGATATGTGT CAGTGCCAAC      60
229 AATGCTACCA CAGTTGCACC TTCTGTAGGA ATTACAAGAT TAATTAAGTC ATCAACGGCA      120
231 GAACCAAGTTA AAGAAGAGGC CAAACTTCA AATCCAACCTT CTTCACTAAC TTCTCTTTCT      180
233 GTGGCACCAA CATTACAGCC AAATATAACT CTGGGACCCA CCTATTTAAC CACTGTCAAT      240
235 TCTTCAGACT CTGACAATGG GACCACAAGA ACAGCAAGCA CCAATTCTAT AGGCATTACA      300
237 ATTCACCAA ATGGAACGTG GCTTCCAGAT AACCAGTTCA CGGATGCCAG AACAGAACCC      360
239 TGGGAGGGGA ATTCCAGCAC CGCAGCAACC ACTCCAGAAA CTTTCCCTCC TTCAGGTAAT      420
241 TCTGACTCGA AGGACAGAAG AGATGAGACA CCAATTATTG CGGTGATGGT GGCCCTGTCC      480
243 TCTCTGCTAG TGATCGTGTG TATTATCATA GTTTTGATCA TGTTAAGGTT TAAGAAATAC      540
245 AAGCAAGCTG GGAGCCATTC CAATTCTTTC CGCTTATCCA ACGGCCGCAC TGAGGATGTG      600
247 GAGCCCCAGA GTGTGCCACT TCTGGCCAGA TCCCCAAGCA CCAACAGGAA ATACCCACCC      660
249 CTGCCCCTGG ACAAGCTGGA AGAGGAAATT AACCAGGAGAA TGGCAGACGA CAATAAGCTC      720
251 TTCAGGGAGG AATTCAACGC TCTCCCTGCA TGTCCCTATCC AGGCCACCTG TGAGGCTGCT      780
253 TCCAAGGAGG AAAACAAGGA AAAAAATCGA TATGTAAACA TCTTGCCTTA TGACCACTCT      840
255 AGAGTCCACC TGACACCGGT TGAAGGGGTT CCAGATTCTG ATTACATCAA TGCTTCATTC      900
257 ATCAACGGTT ACCAAGAAAA GAACAAATTC ATTGCTGCAC AAGGACCAAA AGAAGAAACG      960
259 GTGAATGATT TCTGGCGGAT GATCTGGGAA CAAAACACAG CCACCATCGT CATGGTTACC      1020
261 AACCTGAAGG AGAGAAAGGA GTGCAAGTGC GCCCAGTACT GGCCAGACCA AGGCTGCTGG      1080
263 ACCTATGGGA ATATTCGGGT GTCTGTAGAG GATGTGACTG TCCTGGTGGA CTACACAGTA      1140
265 CGGAAGTTCT GCATCCAGCA GGTGGGCGAC ATGACCAACA GAAAGCCACA GCGCCTCATC      1200
267 ACTCAGTTCC ACTTTACCAG CTGGCCAGAC TTTGGGGTGC CTTTACCCC GATCGGCATG      1260
269 CTCAAGTTCC TCAAGAAGGT GAAGGCTGT AACCCTCAGT ATGCAGGGGC CATCGTGGTC      1320
271 CACTGCAGTG CAGGTGTAGG GCGTACAGGT ACCTTTGTCT TCATTGATGC CATGCTGGAC      1380
273 ATGATGCATA CAGAACGGAA GGTGGACGTG TATGGCTTTG TGAGCCGGAT CCGGGCAGAC      1440
275 CGCTGCCAGA TGGTGCAAAC CGATATGCAG TATGTCTTCA TATACCAAGC CTTTCTGGAG      1500
277 CATTATCTCT ATGGAGATAC AGAACTGGAA GTGACCTCTC TAGAAACCCA CCTGCAGAAA      1560
279 ATTTACAACA AAATCCCAGG GACCAGCAAC AATGGATTAG AGGAGGAGTT TAAGAAGTTA      1620
281 ACATCAATCA AAATCCAGAA TGACAAGATG CGGACTGGAA ACCTTCCAGC CAACATGAAG      1680
283 AAGAACCGTG TTTTACAGAT CATTCCATAT GAATTCAACA GAGTGATCAT TCCAGTTAAG      1740
285 CGGGGCGAAG AGAATACAGA CTATGTGAAC GCATCCTTTA TTGATGGCTA CCGGCAGAAG      1800
287 GACTCCTATA TCGCCAGCCA GGGCCCTCTT CTCCACACAA TTGAGGACTT CTGGCGAATG      1860
289 ATCTGGGAGT GGAAATCCTG CTCTATCGTG ATGCTAACAG AACTGGAGGA GAGAGGCCAG      1920
291 GAGAAGTGTG CCCAGTACTG GCCATCTGAT GGACTGGTGT CCTATGGAGA TATTACAGTG      1980
293 GAACTGAAGA AGGAGGAGGA ATGTGAGAGC TACACCGTCC GAGACCTCCT GGTCACCAAC      2040
295 ACCAGGGAGA ATAAGAGCCG GCAGATCCGG CAGTTCCACT TCCATGGCTG GCCTGAAGTG      2100
297 GGCATCCCCA GTGACGGAAA GGGCATGATC AGCATCATCG CCGCCGTGCA GAAGCAGCAG      2160
299 CAGCAGTCAG GGAACACCCC CATCACCGTG CACTGCAGCG CCGGGGCAGG AAGGACGGGG      2220
301 ACCTTCTGTG CCCTGAGCAC CGTCTGGAG CGTGTGAAAG CAGAGGGGAT TTTGGATGTC      2280

```

## RAW SEQUENCE LISTING

DATE: 10/18/2004

PATENT APPLICATION: US/10/777,144

TIME: 09:18:43

Input Set : N:\Crif3\RULE60\10777144.raw.txt

Output Set: N:\CRF4\10182004\J777144.raw

```

303 TTCCAGACTG TCAAGAGCCT GCGGCTACAG AGGCCACACA TGGTCCAGAC ACTGGAACAG      2340
305 TATGAGTTCT GCTACAAGGT GGTGCAGGAG TATATTGATG CATTCTCAGA TTATGCCAAC      2400
307 TTCAAGTAA      2409
310 (2) INFORMATION FOR SEQ ID NO: 3:
312     (i) SEQUENCE CHARACTERISTICS:
313         (A) LENGTH: 793 amino acids
314         (B) TYPE: amino acid
315         (D) TOPOLOGY: linear
317     (ii) MOLECULE TYPE: protein
319     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
321 Met Asp Ser Trp Phe Ile Leu Val Leu Phe Gly Ser Gly Leu Ile His
322   1          5          10          15
324 Val Ser Ala Asn Asn Ala Thr Thr Val Ser Pro Ser Leu Gly Thr Thr
325          20          25          30
327 Arg Leu Ile Lys Thr Ser Thr Thr Glu Leu Ala Lys Glu Glu Asn Lys
328          35          40          45
330 Thr Ser Asn Ser Thr Ser Ser Val Ile Ser Leu Ser Val Ala Pro Thr
331          50          55          60
333 Phe Ser Pro Asn Leu Thr Leu Glu Pro Thr Tyr Val Thr Thr Val Asn
334          65          70          75          80
336 Ser Ser His Ser Asp Asn Gly Thr Arg Arg Ala Ala Ser Thr Glu Ser
337          85          90          95
339 Gly Gly Thr Thr Ile Ser Pro Asn Gly Ser Trp Leu Ile Glu Asn Gln
340          100         105         110
342 Phe Thr Asp Ala Ile Thr Glu Pro Trp Glu Gly Asn Ser Ser Thr Ala
343          115         120         125
345 Ala Thr Thr Pro Glu Thr Phe Pro Pro Ala Asp Glu Thr Pro Ile Ile
346          130         135         140
348 Ala Val Met Val Ala Leu Ser Ser Leu Leu Val Ile Val Phe Ile Ile
349          145         150         155         160
351 Ile Val Leu Tyr Met Leu Arg Phe Lys Lys Tyr Lys Gln Ala Gly Ser
352          165         170         175
354 His Ser Asn Ser Phe Arg Leu Ser Asn Gly Arg Thr Glu Asp Val Glu
355          180         185         190
357 Pro Gln Ser Val Pro Leu Leu Ala Arg Ser Pro Ser Thr Asn Arg Lys
358          195         200         205
360 Tyr Pro Pro Leu Pro Val Asp Lys Leu Glu Glu Glu Ile Asn Arg Arg
361          210         215         220
363 Met Ala Asp Asp Asn Lys Leu Phe Arg Glu Glu Phe Asn Ala Leu Pro
364          225         230         235         240
366 Ala Cys Pro Ile Gln Ala Thr Cys Glu Ala Ala Ser Lys Glu Glu Asn
367          245         250         255
369 Lys Glu Lys Asn Arg Tyr Val Asn Ile Leu Pro Tyr Asp His Ser Arg
370          260         265         270
372 Val His Leu Thr Pro Val Glu Gly Val Pro Asp Ser Asp Tyr Ile Asn
373          275         280         285
375 Ala Ser Phe Ile Asn Gly Tyr Gln Glu Lys Asn Lys Phe Ile Ala Ala
376          290         295         300
378 Gln Gly Pro Lys Glu Glu Thr Val Asn Asp Phe Trp Arg Met Ile Trp

```

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/10/777,144

DATE: 10/18/2004  
TIME: 09:18:44

Input Set : N:\Crf3\RULE60\10777144.raw.txt  
Output Set: N:\CRF4\10182004\J777144.raw

**Please Note:**

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:9; Xaa Pos.8,12,21,22,24,25,27,28,30,37,39,47,57,72,77,89,94,95,99,104  
Seq#:9; Xaa Pos.109,111,115,116,124,125,131,133,135,137,138,139,143,144,153  
Seq#:9; Xaa Pos.155,170,174,176,179,180,181,182,183,186,205,211,212,214,215  
Seq#:9; Xaa Pos.217,222,227,230,232,240,244,247  
Seq#:14; Xaa Pos.10,20,21,22,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38  
Seq#:14; Xaa Pos.39,40,41,42,43,55,67,70,87,95,96,97,99,100,103,105,106,112  
Seq#:14; Xaa Pos.114,115,116,120,121,123,126,127,128,133,137,138,139,141  
Seq#:14; Xaa Pos.143,149,151,182,186,188,195,196,197,198,199,200,201,202  
Seq#:14; Xaa Pos.205,209,212,213,214,218,222,224,228,229,244,247,254,257  
Seq#:14; Xaa Pos.264,265,266,267

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/777,144

DATE: 10/18/2004

TIME: 09:18:44

Input Set : N:\Crf3\RULE60\10777144.raw.txt

Output Set: N:\CRF4\10182004\J777144.raw

L:28 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]  
L:29 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]  
L:36 M:238 W: Alpha Fields not Ordered, Reordered [(A) APPLICATION NUMBER:] of (1) (vii)  
L:837 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9 after pos.:0  
M:341 Repeated in SeqNo=9  
L:1143 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:14 after pos.:0  
M:341 Repeated in SeqNo=14